

Galaxy clusters in local Universe simulations without density constraints: a long uphill struggle

Jenny G. Sorce^{1,2★}

¹Univ Lyon, Univ Lyon1, Ens de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230, Saint-Genis-Laval, France

²Leibniz-Institut für Astrophysik, An der Sternwarte 16, D-14482 Potsdam, Germany

Accepted 2018 June 16. Received 2018 June 16; in original form 2018 January 9

ABSTRACT

Galaxy clusters are excellent cosmological probes provided that their formation and evolution within the large-scale environment are precisely understood. Therefore, studies with simulated galaxy clusters have flourished. However, detailed comparisons between simulated and observed clusters and their population – the galaxies – are complicated by the diversity of clusters and their surrounding environment. An original way initiated by Bertschinger as early as 1987, to legitimize the one-to-one comparison exercise down to the details, is to produce simulations constrained to resemble the cluster under study within its large-scale environment. Subsequently, several methods have emerged to produce simulations that look like the local Universe. This paper highlights one of these methods and its essential steps to get simulations that not only resemble the local large-scale structure but also that host the local clusters. It includes a new modelling of the radial peculiar velocity uncertainties to remove the observed correlation between the decreases of the simulated cluster masses and of the amount of data used as constraints with the distance from us. This method has the particularity to use solely radial peculiar velocities as constraints: no additional density constraints are required to get local cluster simulacra. The new resulting simulations host dark matter haloes that match the most prominent local clusters such as Coma. Zoom-in simulations of the latter and of a volume larger than the $30 h^{-1}$ Mpc radius inner sphere become now possible to study local clusters and their effects. Mapping the local Sunyaev–Zel’dovich and Sachs–Wolfe effects can follow.

Key words: methods: numerical – techniques: radial velocities – galaxies: clusters: individual – large-scale structure of Universe.

1 INTRODUCTION

Galaxy clusters are unconditional cosmological probes. However, for that purpose their formation and evolution within the large-scale environment must be understood precisely. This task is greatly complicated by the diversity of clusters (Struble & Rood 1988) and their surrounding environment that undoubtedly affect them and their constituents: the galaxies. One-to-one comparisons between simulated and observed clusters and their population as well as maps of clusters’ effects like the Sunyaev–Zel’dovich (Zeldovich & Sunyaev 1969) and the Sachs–Wolfe (Sachs & Wolfe 1967) effects but also gravitational lensing effects obtained numerically must then be considered carefully.

An original way initiated by Bertschinger as early as 1987 (Bertschinger 1987), to legitimize the one-to-one comparison exercise down to the details, is to produce simulations constrained to resemble the object under study. Thus, while typical simulations (Efsthathiou & Eastwood 1981) cope with cosmic variance

by simulating large volumes (e.g. Angulo et al. 2012; Alimi et al. 2012; Vogelsberger et al. 2014; Schaye et al. 2015; Dubois et al. 2016; Dolag, Komatsu & Sunyaev 2016, Euclid’s flagship, for a non-exhaustive list) that abide by a prior cosmological model, constrained simulations consist generally in smaller volumes, abiding also by a cosmological model, but whose initial conditions have been constrained by observational data to decrease the cosmic variance. Many techniques resulting in multiple studies have been developed since then (e.g. Bertschinger 2001; Lavaux 2010; Kitaura 2013; Lavaux 2016; Sorce et al. 2016a; Wang et al. 2016). These methods involve several steps from the treatment via modelling of the observational constraints be they redshift surveys (e.g. Skrutskie et al. 2006; Aihara et al. 2011; Lavaux & Hudson 2011; Huchra et al. 2012) or peculiar velocities (e.g. Mathewson, Ford & Buchhorn 1992; Willick et al. 1997; Zaroubi et al. 2001; Springob et al. 2007; Tully et al. 2008, 2013; Tully, Courtois & Sorce 2016) to the reconstruction of the initial conditions that will result in the object of study. While peculiar velocities constitute more of a challenge for observers than redshifts, the advantage of peculiar velocities over redshift surveys (used as density maps) is that they are correlated on large scales and are highly linear. Constrained initial condi-

* E-mail: jenny.sorce@univ-lyon1.fr

tions can be built either forwardly (e.g. Kitaura & Enßlin 2008; Jasche & Wandelt 2013; Wang et al. 2013) or backwardly (e.g. Bertschinger & Dekel 1989; Dekel, Bertschinger & Faber 1990; Zaroubi, Hoffman & Dekel 1999; Ganon & Hoffman 1993; Lavaux et al. 2008). In the former case, the initial density field is sampled from a probability distribution function consisting of a prior and a likelihood given the observational data. In the latter case, further explained in the next section of this paper, the initial density field is directly obtained from a realization of the density field today.

To summarize, the steps to finally get the product ‘a constrained simulation’ are numerous and can be computer time demanding. It can thus be quite challenging to orientate oneself in the diversity of methods (with all the different steps, assumptions made and computer time they might involve) as well as to understand exactly down to which scale and which area (depending on the survey coverage) these simulations are efficient simulacra of what they are forced to resemble. This knowledge is however essential to lead studies and to draw appropriate conclusions.

This paper investigates one of the backwards methods with all the steps it involves to get simulations that resemble the local Universe, down to the cluster scale, without adding any additional density constraints: the observational information comes from the velocity field only (galaxy radial peculiar velocities), no supplementary assumption is made on the mass of the different local clusters. This paper thus introduces additionally a better modelling of the uncertainties attributed to the peculiar velocities to obtain the final product: a simulation of the local Universe that resemble not only its large-scale structure but that also hosts dark matter haloes sufficiently massive to be considered as local observed cluster counterparts up to at least the distance of the Coma cluster ($70 h^{-1}$ Mpc). Previous studies indeed revealed that several refinements in the modelling of the data and the techniques are required to finally get simulacra of the clusters when using only (i.e. without assumptions on the mass of the clusters) peculiar velocities as constraints. Each one of the refinements added to the method, as an additional step, has so far permitted getting proper (mass, position) and stable (in each simulation) Virgo (Sorce et al. 2016a) and Centaurus (Sorce & Tempel 2018) galaxy cluster simulacra. The better modelling of the uncertainties introduced here permits obtaining simulacra for the Coma cluster that are massive enough and a better stability for the Hydra cluster, as well as a better repartition of the local mass in general in the reproduced local large-scale structure between the different clusters.

This paper thus starts with a brief summary of the different steps of the method used so far to model the data and the involved techniques. Then, it presents the refinement in modelling the uncertainties attributed to the peculiar velocities when inserted in the technique used to reconstruct the local density and velocity fields. Finally, comparisons with previous studies made without this new modelling are proposed before concluding.

2 PRODUCTION OF LOCAL UNIVERSE SIMULACRA

2.1 Method for the first and second generations of constrained initial conditions

This subsection is a reminder of the method so far with a brief description of each step and its utility and impact on the resulting local Universe-like simulations:

(i) Catalogue of constraints: The constraints are radial peculiar velocities of galaxies. They are obtained from a catalogue of galaxy distance moduli (Tully et al. 2013). Using $H_0 = 75.2$ ($= 100 h$) $\text{km s}^{-1} \text{Mpc}^{-1}$ (value given in Tully et al. 2013), this catalogue allows to probe distances as large as $250 h^{-1}$ Mpc but 50 per cent (90 per cent) of the data are within $70 h^{-1}$ Mpc ($160 h^{-1}$ Mpc). It consists of a gathering of more than 8000 galaxy direct distance estimates from several distance estimators: about 88 per cent are from the Tully–Fisher (Tully & Fisher 1977) and Fundamental Plane (Colless et al. 2001) relations. The rest is from Cepheids (Freedman et al. 2001), Tip of the Red Giant Branch (Lee, Freedman & Madore 1993), Surface Brightness Fluctuation (Tonry et al. 2001), supernovae of Type Ia (Jha, Riess & Kirshner 2007) and other miscellaneous methods.

(ii) Grouping of constraints and derivation of peculiar velocities: The quality of the reconstruction obtained with a *linear* minimum variance estimator depends strongly on the grouping of individual data points into single points to suppress virial motions in high-density regions (Sorce, Hoffman & Gottlöber 2017b). The grouping must remove all the virial motions without affecting neither galaxies in the field (Sorce et al. 2017b) nor regions of infall (Sorce & Tempel 2017). Only then the proper masses are recovered for the most nearby clusters such as Centaurus (Sorce & Tempel 2018), where the proliferation of distance measurements in these regions might lead to an ‘overgrouping’ making galaxies falling onto the clusters too sparse.

The cluster masses could be artificially increased by adding point-like density constraints but (1) mass estimates of observed clusters are not easily determined, (2) masses would not be recovered with the sole velocity field, consequently statistical studies would be biased by the chosen masses, (3) only selected clusters would have their masses increased. Combining both redshift surveys as density and velocities is challenging: since these two types of surveys are not subject to the same biases, having both surveys predict exactly the same field on a megaparsec scale is difficult. It is then more interesting to focus on using only one type of surveys or eventually both but at different distances: velocities nearby and redshift surveys further away where peculiar velocity measurements become sparse. This paper uses only peculiar velocities directly derived from the grouped distances (of groups). The latter are weighted by their uncertainties linked to the distance estimators used to derive them [see equations 1 to 3 in Sorce & Tempel (2017) and Tully et al. (2013) for detailed explanations on the attribution of uncertainties]. Uncertainties on grouped distances are subsequently propagated to derive peculiar velocity uncertainty. This additional parameter, called σ_{NL} , is determined so that $\chi^2 = c_i \langle c_i c_j \rangle^{-1} c_j$ (quantity to gauge the linearity of the data with c_i and c_j the constraints plus their uncertainties) divided by the number of degrees of freedom (dof, number of constraints) is about 1. When $\chi^2/\text{dof} \sim 1$, the data are then sampling a typical realization of the assumed prior power spectrum model in the linear theory.

(iii) Minimization of biases: The grouping described above is inefficient if used without a bias minimization scheme (Sorce 2015) that prevents the spurious infall onto the local volume, removes the dependence of the reconstruction on the choice of the local H_0 (Sorce & Tempel 2017) and that permits reconstructing sharper structures. This technique also enabled to produce the first simulations that host stable (position, mass, merging history) and massive (more than $2 \times 10^{14} h^{-1} M_\odot$) Virgo dark matter haloes without density constraints (see discussion in (ii) about these density constraints) and with the ‘aggressive’ grouping released with the catalogue of distances (Sorce et al. 2016a,b). It also permitted devel-

oping the local Group factory (Carlesi et al. 2016) at the origin of multiple projects. This minimization scheme relies on the theoretical expectation that the distribution of radial peculiar velocities must be a Gaussian (Sheth & Diaferio 2001). It iteratively reduces spurious non-Gaussianities in the radial peculiar velocity distribution, to retroactively derive overall better distance estimates resulting in a minimization of the effects of biases. Uncertainties σ attributed to the biased minimized distances (velocities) are of 5 per cent (Sorce 2015).

(iv) Relocation and replacement of constraints: Constraints observed today must be relocated to the position of their precursors to get proper initial conditions. At redshift zero in the simulation, objects can then be expected to be at the proper observed positions. The cosmic displacement field is derived with the Wiener-Filter technique (a linear minimum variance estimator of the data assuming a prior cosmological model, Zaroubi et al. 1995, 1999). The constraints are then displaced using the Reverse Zel’dovich Approximation (Doumler et al. 2013).

Additionally, after their relocation, the *noisy and uncertain* radial peculiar velocities must be replaced by their 3D Wiener-Filter reconstructions (Sorce et al. 2014) to ensure that not only do simulations reproduce the local large-scale structure but that there exists stable knots (clear overdensities at clusters’ locations) in each simulation.

(v) Building of initial conditions: peculiar velocity constraints are finally inserted in the constrained realization technique (Hoffman & Ribak 1991, 1992; van de Weygaert & Bertschinger 1996) to derive an estimate of the residual between the Wiener-Filter reconstructed field and the model. It restores statistically the missing structures in the reconstructed field (the Wiener Filter goes to the mean field in absence of data or in presence of very noisy data). It relies on the combination of the constraints and a random realization field.

The resolution of the initial conditions (obtained after rescaling the fields obtained with the constrained realization technique) is increased by adding small-scale features with codes like GINNUNGAGAP¹ and MUSIC (Hahn & Abel 2011). These codes also permit producing zoom initial conditions (Bertschinger 2001) to study with a high-resolution local objects. It consists in keeping the large-scale environment at low resolution for its effect on the region of interest while the resolution is increased solely in this region.

The first and second generations of simulations constrained with the observational data from Tully et al. (2013) were obtained with these steps. While the first generation was based on the groups from the literature released with the catalogue, the second generation used groups obtained with a Friends of Friends (FoF) percolation method, where different linking lengths in radial (along the line of sight) and in transversal (in the plane of the sky) directions are used. Additionally, the groups are refined using multimodality analysis to separate nearby/merging systems (Tempel et al. 2016). The simulations resemble the local large-scale structure (Sorce et al. 2016a) and host Virgo (first, Sorce et al. 2016b) and Centaurus simulacra (second Sorce & Tempel 2018).

The bias minimization was essential in getting such results. Without it, the second generation would produce simulations where the success rate in hosting a Virgo simulacrum is only 60 per cent against 100 per cent and halo masses would barely reach $10^{14} h^{-1} M_{\odot}$. As for Centaurus, the majority of the halo masses

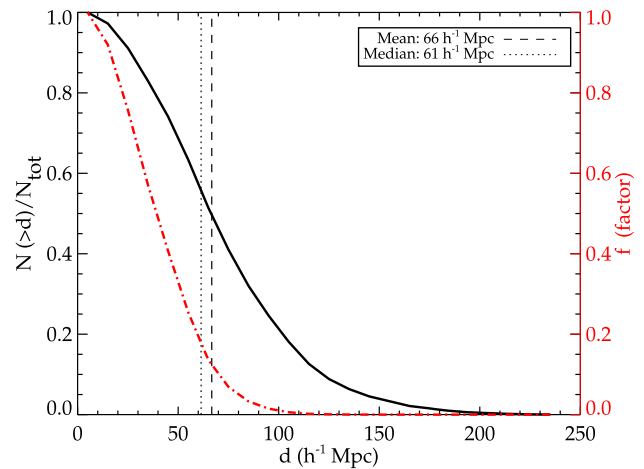


Figure 1. Normalized cumulative distribution of distances (black line) of the catalogue of constraints. The dot-dashed red line is the same distribution to the power 3. It is used as a factor when improving the modelling of the peculiar velocity uncertainties. Dotted and dashed black lines are the median and mean distances of the catalogue.

would barely be above $10^{14} h^{-1} M_{\odot}$ and only a few would reach $4 \times 10^{14} h^{-1} M_{\odot}$, still below recent estimates.

In addition, although a dark matter halo simulacrum is present in most if not all the simulations for local clusters further away than these two very nearby clusters. Their masses are barely above $1.5 \times 10^{14} h^{-1} M_{\odot}$. This is certainly in agreement with observational estimates for clusters like Hydra but not for the Coma cluster.

As discussed previously, there are severe disadvantages in increasing locally the masses of these clusters. Consistency in using the sole peculiar velocities as constraints is thus highly preferable to preserve the interest in using these simulations for statistical purposes.

2.2 Modelling the uncertainties: third generation of constrained initial conditions

It is in a sense completely expected that the accuracy of the simulations drops with the distance from us: there are less data and they are more noisy and uncertain. However, this systematic shrink of the most massive haloes inherent to the technique should be accounted for to produce simulations reproducing the local clusters up to greater distances than those within the first 30 h⁻¹ Mpc and above all to ensure a proper mass distribution within the local volume. A feature of the Wiener Filter and by extension of the Constrained Realization technique is to smooth the data when they are sparser and noisier. While this is absolutely correct when studying the sole reconstruction, the effect of smoothing when there are less data is undesirable for constrained simulations. Surely, the Constrained Realization technique compensates the effect by adding a random realization but unless this realization matches in a sense the observational data by imposing the same type of constraints at a given location, the chances of getting the clusters are drastically reduced with the distance from us. Indeed, the amount of data in the catalogue of constraints drops drastically with the distance (see Fig. 1).

Since the Wiener Filter also smooths the data with respect to their uncertainties plus σ_{NL} added in quadrature (see Section 2.1), it is possible to model differently this error term to account for the smoothing with the distance. The error term – uncertainties σ

¹<https://github.com/ginnungagapgroup/ginnungagap>

(5 per cent after applying the bias minimization scheme) plus σ_{NL} to sample a typical realization of the prior cosmological model – becomes after adding a third term:

$$\sigma'(d) = \sqrt{\sigma(d)^2 + \sigma_{\text{NL}}^2 + (f(d) \times \sigma_{\text{eff}})^2}, \quad (1)$$

where f is the red line factor given in Fig. 1, it depends on the distance d of the constraint whatever the direction of observation. This factor is the normalized cumulative distribution of distances in all directions to the power 3.

By default, σ_{eff} is chosen to be equal to σ_{NL} required to obtain $\chi^2/\text{dof} \sim 1$ before adding the third term. σ'_{NL} can then be determined to obtain $\chi^2/\text{dof} \sim 1$ after adding the third term. It appears that σ'_{NL} is not required anymore: $\sigma'_{\text{NL}} = 0$ gives $\chi^2/\text{dof} \sim 1$. Note that alternatively, σ_{eff} could be adjusted to obtain $\chi^2/\text{dof} \sim 1$ so that in any case σ'_{NL} is not required anymore. Since there is no need to adjust σ_{eff} here, this alternative is not further investigated. The modelling of the uncertainties as proposed above is thus similar to replacing $\sqrt{\sigma^2 + \sigma_{\text{NL}}^2}$ by $\sqrt{\sigma^2 + (f \times \sigma_{\text{eff}})^2}$, where σ_{NL} used to be applied uniformly to the whole dataset. The purpose was to filter out the non-linear residual in the catalogue of constraints but it did not take into account the filtering because of the decrease in the amount of data with the distance from us. Interestingly, required values of σ_{eff} and σ_{NL} necessary to obtain $\chi^2/\text{dof} \sim 1$ are identical.

This σ_{eff} is expected to force the Wiener Filter to smooth more the nearby data points than those further away to balance the opposite smoothing due to the decrease of the number of constraints with the distance. It does not affect the smoothing due to the noise and uncertainties directly linked to the distance (peculiar velocity) estimates.

One could argue that the number of data point is not isotropically fading with the distance. Still, Sorce et al. (2017a) showed that the Wiener Filter is smoothing almost isotropically for the catalogue used here. Adding this function of space in the factor f does not improve the results presented in the next section for distant local clusters and degrades the very nearby clusters suggesting that this extra complexity does not constitute a solution to the weak anisotropy. Indeed, small fluctuations of σ in the small inner volume because of a small anisotropy is not optimal for the Wiener Filter that works best on the large scales. In light of the above and since there is a priori no reason to use the anisotropic function on the large distance and not close by and since this anisotropy is almost inexistent, it is disregarded in the rest of the paper.

Subsequently, some tests were conducted on the mock catalogues developed in Sorce (2015) to further support this new modelling of the uncertainties. These mock catalogues, described in details in Sorce (2015), reproduce the observational biases and the distribution of grouped data in the catalogue of peculiar velocities. Before the modelling of uncertainties as proposed above, mass functions of re-simulations of the reference simulation obtained with the mock catalogues differ from that of the reference simulation at the large mass end. The reference simulation has on average 30 per cent more haloes of masses above $10^{14} h^{-1} M_{\odot}$ than the re-simulations. In other words, the re-simulations lack 30 per cent of massive haloes on average. This proves that the smaller mass functions at the high mass end found for the constrained simulations are not entirely due to the cosmic variance. On the opposite, the uncertainty modelling described above permits recovering on average the mass function of the reference simulation at the large mass end in the new re-simulations: on average, the re-simulations have the same number of haloes of masses greater than $10^{14} h^{-1} M_{\odot}$ as the reference simulation. Such tests are in favour of modelling the uncertainties

according to the fading of the constraints with the distance. Furthermore, any residual difference between the constrained simulations and the random simulations will then be a result of the cosmic variance.

The initial conditions for the third generation of simulations constrained with the radial peculiar velocity catalogue are then obtained in the standard way. Basically, the steps are those described above (i to v) but in step (iii) the uncertainties σ are replaced with those given by equation (1).

3 COMPARISONS BETWEEN THE THREE GENERATIONS

Three sets of nine simulations each sharing the same nine random realization (in the constrained realization technique) are further used for the comparisons in that section. These sets are called first, second and third generations. The difference between the first and second generations has been described above in detail. It resides in the grouping scheme. The third generation differ from the second by the modelling of the uncertainties after the minimization scheme. The importance of this scheme has also already been discussed. Table 1 summarizes the different steps for the three generations as well as their relative importance.

All these simulations are $500 h^{-1} \text{Mpc}$ boxes run with 512^3 particles, a resolution sufficient to study clusters at redshift zero (particle mass about $8 \times 10^{10} h^{-1} M_{\odot}$, namely about 10^3 or more particles in objects larger than $10^{14} h^{-1} M_{\odot}$). They are run within the framework of Planck cosmology ($\Omega = 0.307$, $\Omega_{\Lambda} = 0.693$, $H_0 = 67.77 \text{ km s}^{-1} \text{Mpc}^{-1}$, $\sigma_8 = 0.829$, Planck Collaboration XVI 2014).

To perfect the comparisons, a set of nine random simulations (sharing the same random realizations and settings) are also run.

3.1 The local large-scale structure

Fig. 2 shows the supergalactic XY slice of one simulation from each set of the three generations. The simulations share the same random realization. Solid (dashed) contours stand for the overdensities (underdensities). The green colour delimits the overdensities from the underdensities. A few structures are identified with blue names. Overall the local large-scale structure is similar in the three cases confirming the efficiency of the constraining scheme on the large scales. However, differences appear at the cluster scale: from left to right, clearly the Coma region is getting denser and denser. The same goes for the Shapley concentration on the left side of Coma. Perseus is also a bit sharper from left to right. Only the densities for Virgo and Centaurus regions increase to decrease again. Since the masses of Virgo and Centaurus in the second generation are slightly larger than observational estimates (Sorce & Tempel 2018), this is not in conflict with observations in a first qualitative approach. A detailed study of the masses in the next subsection confirms this assertion in a quantitative way.

To enhance these qualitative observations, cell-to-cell comparisons between pairs of simulations are conducted. First, cells are compared within the full box. The scatter around the 1:1 relation is derived. Once all the scatters are obtained for a given type of simulation pairs, their mean and variance are computed. Secondly, because simulations are known to be more constrained in the centre of the box where most of the constraints are, cells are compared only in sub-boxes. All the resulting mean scatters (as defined above when comparing the full boxes and different size sub-boxes) and their variances are reported in Fig. 3 as a function of the size of the sub-boxes within which cells are compared between simulations.

Table 1. Overview of the three generations of constrained simulations. The steps of the method are reminded and the differences between the three generations are highlighted. The last line gives the importance or impact of the step on the resulting dark matter halo simulacra. The local large-scale structure is simulated in all the cases.

Generation	Steps in the method					
	(i) Catalogue (Tully et al. 2013)	(ii) Grouping	(iii) Bias Minimization	(iv) a Reverse Zel'dovich (Doumler et al. 2013)	(iv) b 3D (Sorce et al. 2014)	(v) Constrained realizations (Hoffman & Ribak 1991)
First (Sorce et al. 2016a)	✓	Tully et al. 2013	Sorce 2015	✓	✓	✓
Second (Sorce & Tempel 2018)	✓	Sorce & Tempel 2017	Sorce 2015	✓	✓	✓
Third (This paper)	✓	Sorce & Tempel 2017	Sorce 2015 + uncertainty modelling (this paper)	✓	✓	✓
Importance		Nearby masses, stability + farther masses		Positions	Stability, knots	

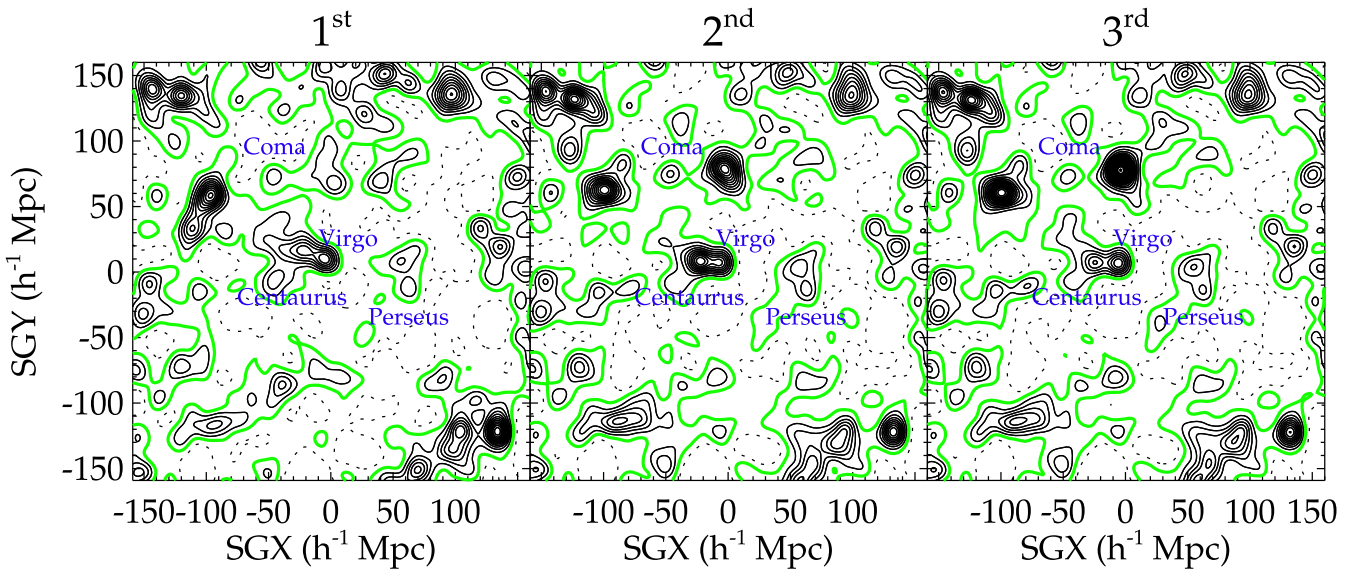


Figure 2. XY supergalactic slices of the local large-scale structure obtained in three constrained simulations sharing the same random realization. Different grouping schemes are used to remove non-linear motions from the constraint-catalogue of galaxy radial peculiar velocities and a new modelling of the uncertainties permits accounting for the smoothing due to the decrease in the number of constraints with the distance. Left: grouping scheme released with the catalogue. Middle: grouping scheme based on an advanced friend-of-friend algorithm. Right: same grouping scheme but with the new uncertainty modelling presented in this paper. In all three cases, the bias minimization scheme is applied. The contours stand for the density field. Solid lines show overdensities while dashed lines represent underdensities. The green colour is the mean field. A few structures are named in blue. Overall the local large-scale structure is properly reproduced in the three simulations but differences appear at the cluster scale. For instance, the overdensity at Coma's location is sharper and sharper from left to right.

The left panel of the figure gives the cosmic variance of the three generations of constrained simulations ($C1^{st}$, $C2^{nd}$, $C3^{rd}$) as well as that of the random simulations (R). As already shown for the two first generations (Sorce et al. 2016a; Sorce & Tempel 2018), the cosmic variance is decreased by a factor 2 to 3 within the inner part of the box: the scatters obtained when comparing pairs of constrained simulations of the same nature are considerably smaller than those obtained when comparing random simulations. Interestingly, for the third generation of constrained simulations presented in this paper, the cosmic variance in the inner part of the box is even smaller than for the two other generations suggesting a slightly better control of the constrained region for that last generation.

The two other panels of the figure show the cosmic variance between simulations of different types be they constrained and random or constrained but of different generations. The comparisons between the first and second generations has already been led by Sorce & Tempel (2018). They are thus not reiterated here. In the middle panel, mean scatters and their variances are derived only with pairs of simulations that do not share the same random realization. On the contrary, the right panel gives the mean scatters and their variances only for pairs of simulations sharing the same random realization.

Comparisons between pairs of simulations that do not share the same random realization (middle panel) show that overall constrained simulations differ in a similar way from the random sim-

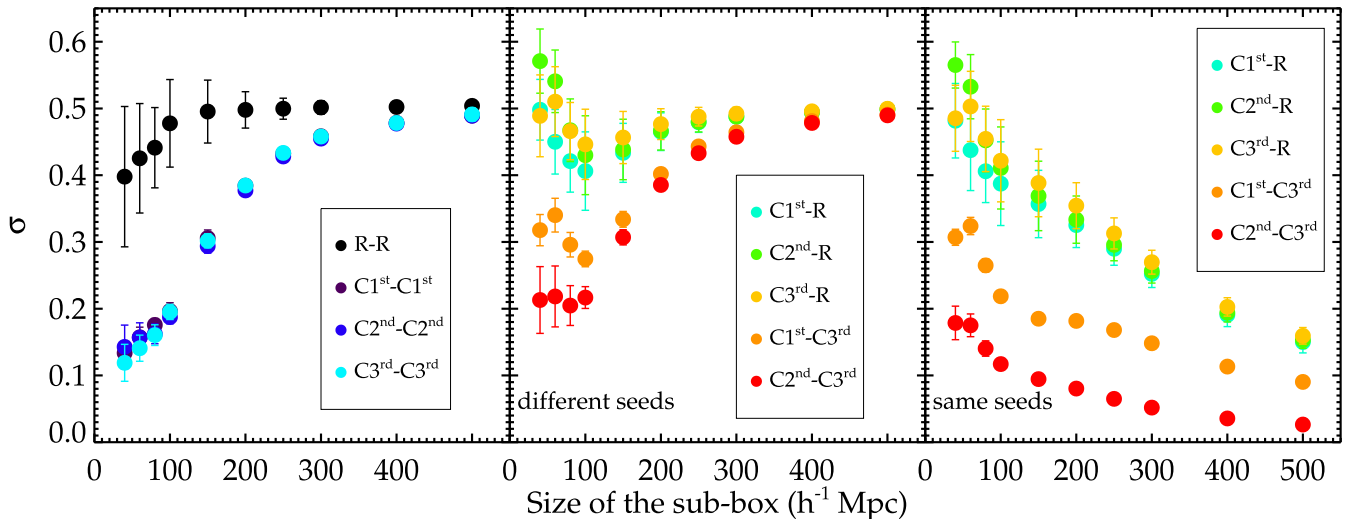


Figure 3. Average variance (filled circle) and its standard deviation (error bar) between density fields of simulations as a function of the size of the compared sub-box. From left to right: comparisons of pairs of random (R, black) and constrained ($C1^{st}$ violet, $C2^{nd}$ dark blue, $C3^{rd}$ light blue) simulations, comparisons between random and constrained simulations (light blue, green and yellow) as well as between constrained simulations obtained with different groupings or uncertainty modellings (orange and red) that do not share the same random realization (middle) and that share the same random realization (right). The goal of the constrained simulations is fulfilled: the cosmic variance is reduced with respect to that of random simulations.

ulations (light blue, green and yellow). There are small differences in the inner part of the box where the second and third generation simulations differ more from the random simulations than those of the first generation. Interestingly, the second generation simulations versus the random simulations present the highest scatter in the inner $50 h^{-1}$ Mpc, while the third generation simulations versus the random ones do so at larger distances. Within the inner box ($30 h^{-1}$ Mpc) the latter have the same scatter as the first versus random simulations. It is a first hint at a better repartition of the mass in the local large-scale structure between the different clusters for the third generation simulations. Indeed, the second generation simulations have very nearby clusters slightly overmassive while further clusters are undermassive. This can explain the larger scatters observed in the very central part of the box: larger than expected overdensities to be compared to overall voids in random simulations – higher probability of finding a void in a small sub-box in a random realization than in a constrained simulation since structures are there by construction. The first generation simulations have undermassive clusters everywhere except the Virgo cluster that is in good agreement with observations. Hence the smaller scatters. As for the third generation simulations versus the random simulations, since the mass is better distributed in the constrained simulations, the scatters are overall of the same order in all the sub-boxes. Comparisons between the different generation simulations confirm the impact of both the grouping scheme and the new modelling of the uncertainty in the inner part of the box: beyond $160 h^{-1}$ Mpc, the measured scatters are similar as those obtained when comparing simulations of the same generation (left panel). However, within the $160 h^{-1}$ Mpc, the scatters are larger when comparing simulations of different generations (middle panel) than those of the same generation (left panel).

Finally, the right panel gives the variance between pairs of simulations that share the same random realization. As expected, when considering the full box, since the non-constrained regions overcome the constrained ones, the scatters are the smallest. They then increase when considering smaller and smaller sub-boxes. In other words, the random realization dominates in the whole volume thus

comparisons involve almost the same fields. However, within the inner $160 h^{-1}$ Mpc part of the box, the scatters have the same values as those obtained when comparing simulations that do not share the same random realization. Overall the scatters follow the same order: they increase from the comparisons between the simulations of the second and third generations to the comparisons between the third generation and random simulations. The same conclusions as before can thus be drawn.

Another quantitative analysis can be led by comparing the power spectra and the mass functions of all these simulations. To derive the mass functions, the dark matter haloes are identified with Amiga's halo finder (Knollmann & Knebe 2009). Fig. 4 presents the ratios of all these power spectra (top panel) and mass functions for the entire boxes (middle panel) and for the $160 h^{-1}$ Mpc radius spheres (bottom panel). The first panel of the figure shows that the power spectra of the constrained simulations are overall below those of the random simulations (mean of the yellow and blue areas below 1) as already observed by Sorce et al. (2016a). However, the second and third generation simulations have slightly larger power spectra. Mostly on the large scales ($> 100 h^{-1}$ Mpc) for the second generation simulations while it is better distributed on all scales above about $30 h^{-1}$ Mpc for the third generation (mean of the red area above 1 for $k < 0.03 h^{-1}$ Mpc). This is overall a good result since the probability of having such a power spectrum given a prior (Planck) for the local Universe is increased from the first to the third generation.

Regarding the mass functions, the same observations can be made. Clearly the second and third generation simulations have more massive haloes than the first generation ones. This affirmation is more visible within the $160 h^{-1}$ Mpc radius sphere. The red area has clearly a mean above 1 at the high mass end. Namely, the third generation simulations have more massive haloes than the first generation simulations. With respect to the second generation simulations, the more massive haloes are better distributed over the whole high mass end in the third generation simulations: the mean green area is slightly above 1 for masses greater than $10^{14} h^{-1} M_{\odot}$ but below 1 at the extreme end ($\sim 10^{15} h^{-1} M_{\odot}$). These observations

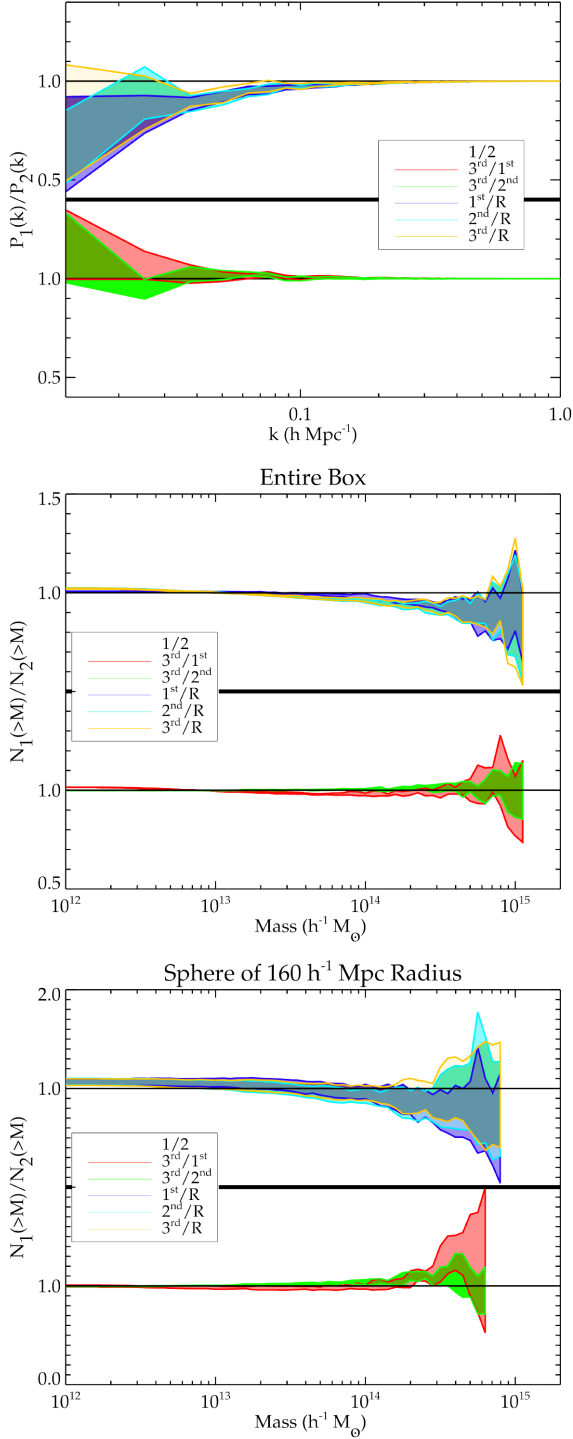


Figure 4. Top: 1σ confidence interval of the ratio of the power spectra of constrained and random simulations (blue and yellow) and of constrained simulations (red and green). Middle and bottom: the same as the top panel but for the mass functions of the entire box and of a $160 h^{-1}$ Mpc radius sphere.

are again in favour of a better distribution of the mass within the local volume.

To conclude this subsection, the local large-scale structure is over-all reproduced by the simulations in all three generations. However, the second then third generation simulations have more probable power spectra with respect to Planck's. Additionally they have more

massive haloes. The third generation simulations are the most probable given the prior cosmological model and have a better repartition of the mass within the local clusters. The next subsection will shade some lights on this assertion by studying in detail the haloes found in the simulations.

3.2 Local galaxy clusters

Simulacra of five of the local clusters (Virgo, Centaurus, Hydra, Coma and Perseus) are looked for in the three sets of constrained simulations. A dark matter halo is considered as a simulacrum of a local observed cluster if it is at the proper location with respect to observational estimates (observed and simulated distances cannot differ by more than 30 per cent) and if its mass is above $10^{14} h^{-1} M_{\odot}$. It is important to note that if simulacra are slightly shifted in positions with respect to the observed cluster, their shifts are consistent between each other. Namely simulacra are only a few megaparsecs away from one simulation to another. The top row of Fig. 5 gives the probability of finding a simulacrum for a given cluster in each set of constrained simulations. A unique Virgo (Centaurus) simulacrum is found in each one of the simulation, whatever the generation is. A Coma look-alike is however found in only 65 per cent of the first generation simulations while it is in a 100 per cent of the other generation simulations. The probability of finding a Hydra simulacra is increased up to 90 per cent in the third generation simulations compared to the 80 per cent before. Perseus is the only local cluster whose probability decreases in the second and third generations. It is worth noting that this is only due to the cut in mass at $10^{14} h^{-1} M_{\odot}$. Would the cut be slightly smaller then the probability in the second and third generation simulations would increase since the halo candidates that are rejected are barely below $10^{14} h^{-1} M_{\odot}$ in these cases. In addition, the currently used catalogue of constraints does not have much data points in that region. With the third catalogue of constraints that contains more distances in that region (Tully et al. 2016), better results can be expected for that particular area. Globally, the third generation of simulations increases the probability of finding simulacra of the local clusters.

The bottom row of that same figure is even more important as it shows the mean, standard deviation, minimum and maximum masses, M_{200} (i.e. the mass enclosed in a sphere with a mean density of 200 times the critical density of the Universe), of these dark matter simulacra. Comparing these M_{200} to observational estimates is not direct. For instance, Tully (2015) published recently the virial masses of these local clusters in M_{\odot} to be compared with M_{200} . These latter are proportional to the virial mass (given by the halo finder) via a factor of 0.80 ± 0.03 (e.g. Sorce et al. 2016b). Assuming the virial masses given by both observational estimates (cf. Table 2) and the halo finder to be roughly similar, the former are shown for comparisons on Fig. 5 with thick blue dashed lines as well as the 1σ uncertainty of the conversion factor with thinner lines. The red dotted lines stand for a very recent estimate of Virgo's mass (M) via the first turn around radius by Shaya et al. (2017) with $M_{200} = 0.70 \pm 0.10 M$ (Sorce et al. 2016b). Again, only the uncertainties on the conversion factor are retained for the plot.

Virgo and Centaurus dark matter haloes are more massive than in the first generation simulations but less massive than in the second. This is in excellent agreement with observational estimates shown as red and/or blue lines. While masses of Virgo simulacra are compatible with the observational estimates for all three generations within 3σ s, they are within 2σ s only for the first and third generations. As for Centaurus like haloes clearly undermassive in the first generation, their average mass is now clearly on top of the observa-

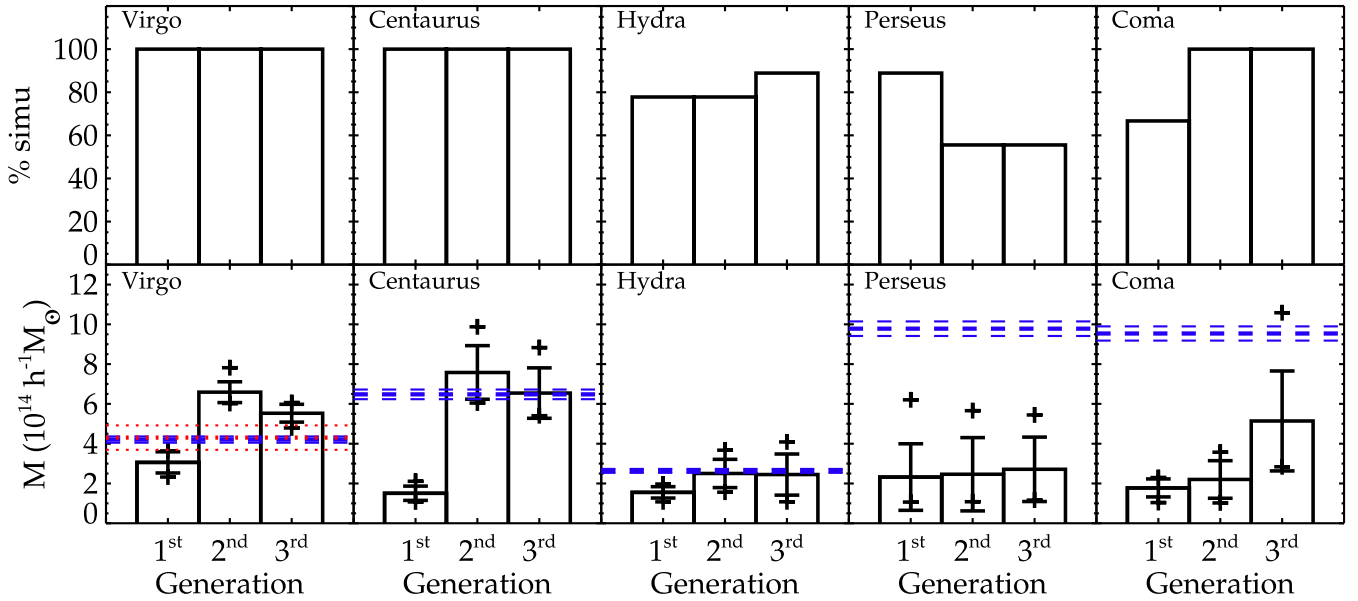


Figure 5. Top row: percentage of simulations with a simulacrum of the observed cluster whose name is given in the top right corner of each panel for the three generations of constrained simulations. Bottom row: average mass (histogram), standard deviation (error bar) and minimum/maximum masses (crosses) of the different simulacra for the three generations. Thick blue dashed and red dotted lines are observational mass estimates. The thinner dashed and dotted lines stand for the uncertainty on the proportionality factor between observational masses and M_{200} . Without any doubt, the third generation simulations are those that host overall the best simulacra of these local clusters.

Table 2. Clusters from Tully (2015) with $H_0 = 75 \text{ km s}^{-1} \text{ Mpc}$: (1) cluster name, (2) supergalactic longitude, (3) supergalactic latitude, (4) distance, (5) virial mass.

(1) Cluster	(2) sgl ($^{\circ}$)	(3) sgb ($^{\circ}$)	(4) d (Mpc)	(5) M ($10^{14} M_{\odot}$)
Virgo	103.0008	− 2.3248	14.9	7.01
Centaurus	156.2336	− 11.5868	38.7	10.8
Hydra	139.4478	− 37.6063	41.0	4.39
Perseus	347.7159	− 14.0594	52.8	16.3
Coma	89.6226	8.1461	73.3	15.9

tional estimates. The slightly smaller masses in the third generation with respect to those in the second generation are even in better agreement with observations. Coma simulacra have more than doubled (almost tripled) their masses. The third generation simulations are the only ones hosting dark matter haloes that are within 2σ s of the observational estimate. Hydra simulacra are in perfect agreement with observations for both the second and third generation simulations. Although this agreement is smaller, the first generation simulations’ masses are still within 3σ s. Perseus-like haloes are the only ones that are not drastically improved. Although the mean mass increases from the first to the third generation simulations, they are still barely within 3σ s and more within 4σ s. As explained before, this region is not well probed by the catalogue of constraints used here.

Regardless, without any doubt, the third generation simulations are those that host overall the best simulacra of the local clusters. The drastic increase in mass of the very nearby cluster-like at the expense of the farther local clusters visible in the second generation simulations is erased by the new modelling of the uncertainties.

The top (bottom) row of Fig. 6 probes the relative change (variation) between simulacra hosted by different (the same) generation

simulations. Virgo look-alike is incredibly stable between the different generations (top row) and within a given generation (bottom row). The only notable difference is in the z direction. This direction is more constrained in the second and third generation simulations than in the first generation ones. This is remarkable since this is the less constrained direction because of the zone of avoidance due to the Milky Way’s dust. Centaurus simulacrum becomes also quite stable. The only exceptions being the y and z components of the velocities. Still the y component is much more stable in the second and third generations than it was before. Dark matter haloes that are substitutes of Coma and Hydra are more stables in terms of positions in the second and third generation simulations. As for Perseus-like, its velocity components are more stable for the second and third generation simulations, especially the x component. Remarkably, its position is more stable only in the third generation simulations, especially the y component.

To conclude the biggest impact of the new modelling of the uncertainties is to reach better agreement between observed and simulated masses.

4 CONCLUSION

This paper investigates the method used so far to get local Universe like simulations that resemble the local Universe down to the cluster scales using only peculiar velocities (no densities) as constraints. Namely not only do these simulations look like the local large-scale structure but they also host dark matter haloes that can be taken as simulacra of local clusters. So far the simulations obtained with the method described in this paper from the catalogue of distances to the constrained realization technique going through a series of steps all more important than each other – grouping, minimization of biases, reverse Zel’dovich approximation, replacement of radial by 3D velocities – permitted getting great simulacra for the Virgo and Centaurus galaxy clusters. Local clusters further away seemed

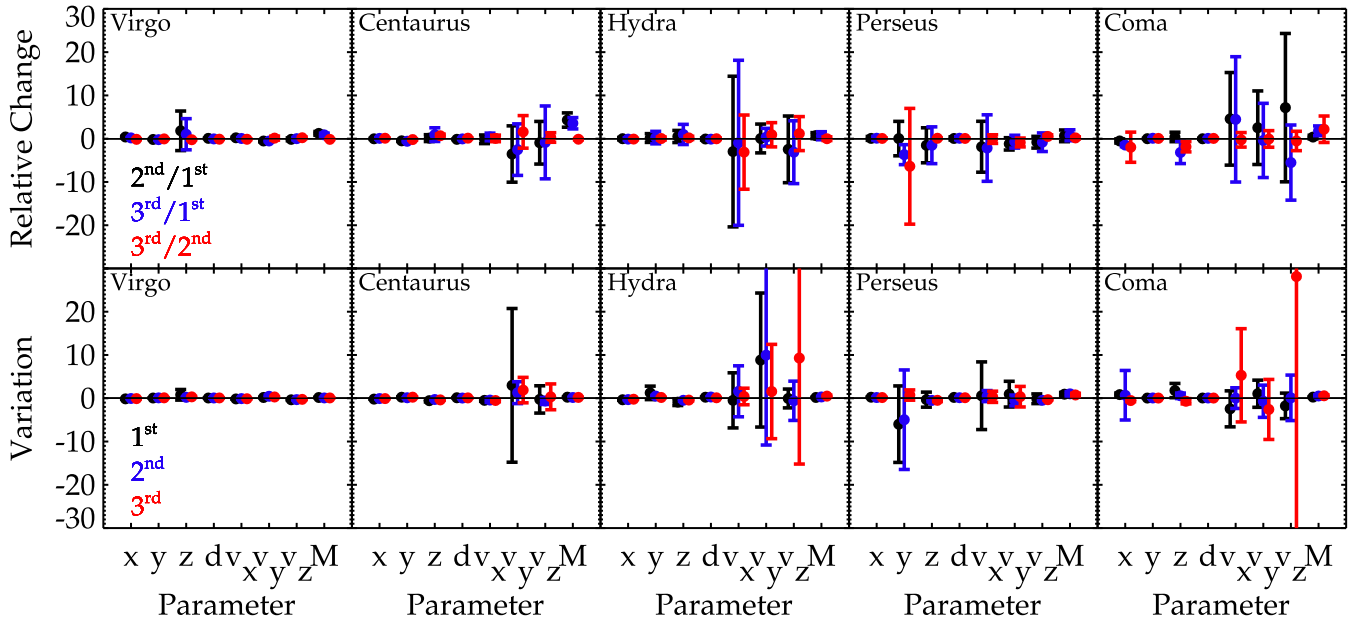


Figure 6. Top row: relative change and standard deviation (filled circle and error bar) between the parameters of the cluster simulacra obtained in the simulations produced with the different grouping schemes or uncertainty modelling. Bottom row: variation (filled circle) and standard deviation (error bar) of the parameters of the cluster simulacra found in the first (black), second (blue) and third (red) generations of constrained simulations.

irremediably too faint without adding any additional density constraints.

However, point-like density constraints are not optimal: (1) mass estimates of clusters are not easily obtained and subject to uncertainties; (2) chosen mass values can bias the statistical results obtained with the resulting constrained simulations. Simulations constrained with both types of constraints cannot cover as much ground as simulations constrained solely with peculiar velocities. For instance, the mass of a cluster induced by a given environment is not deductible anymore since it has been independently constrained; (3) it is impossible to add a density constraint for all the clusters especially for those in the zone of avoidance; (4) adding a full redshift survey to the peculiar velocities is complicated by the fact that both surveys must be greatly consistent to predict the same field (positions included) down to the cluster scale. In light of the above, it seems grandly recommended to simulate well-known observed clusters using solely radial peculiar velocities not to bias the results and conclusions for these clusters but also for other clusters in the local volume.

Observing that the fading of the cluster masses in the constrained region of the simulations is correlated to the decrease in the amount of data in the catalogue with the distance from us (the centre of the box), a new modelling of the peculiar velocity uncertainties after the minimization scheme is proposed. This new modelling is built to account for the smoothing of the reconstruction technique due to the decreasing amount of data with the distance. It does not affect the filtering of the data due to their uncertainties inherent to their measurements. This new modelling appears similar to replacing the uncertainty by a sum in quadrature of this same uncertainty and a factor f times a σ effective. The factor f depends on the distance of the constraint from the observer. It is obtained with the normalized distribution function of the constraints with the distance to the power 3. The σ effective is set by default to the non-linear σ_{NL} value required, before adding σ effective, such that the data are sampling a typical realization of the assumed prior power spectrum

model in the linear theory ($\chi^2/\text{dof} \sim 1$). After adding σ effective, σ_{NL} appears not to be required anymore ($\sigma_{NL} = 0$).

Multiple comparisons between the previous generation simulations to those obtained with this new modelling of the uncertainties reveal that the latter have overall a power spectrum with a higher probability given the prior cosmological model (Planck) and that they have a better distribution of the mass within the local large-scale structure between the different clusters. Consequently, although both the previous and new generation simulations overall reproduce the local large-scale structure, the latter generation permits producing simulacra of local clusters further away than Virgo and Centaurus such as the Coma cluster with appropriate masses.

Zoom-in simulations of the Coma cluster and of a larger volume than the $30 h^{-1}$ Mpc radius inner sphere become now possible to study local clusters and their effects. Mapping the impact of the local Universe on distant observations can follow.

ACKNOWLEDGEMENTS

JS acknowledges stimulating discussions with her collaborators: Klaus Dolag, Stefan Gottlöber and Yehuda Hoffman. The author would like to thank the referee for comments that helped improve the manuscript. JS acknowledges support from the Centre National d'études spatiales (CNES) postdoctoral fellowship program as well as from the 'l'Oréal-UNESCO Pour les femmes et la Science' fellowship program. The author gratefully acknowledges the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for providing computing time on the GCS Supercomputer SuperMUC at LRZ Munich.

REFERENCES

- Aihara H. et al., 2011, *ApJS*, 193, 29
- Alimi J.-M. et al., 2012, preprint (arXiv: [1206.2838](https://arxiv.org/abs/1206.2838))
- Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, *MNRAS*, 426, 2046

- Bertschinger E., 1987, *ApJ*, 323, L103
- Bertschinger E., 2001, *ApJS*, 137, 1
- Bertschinger E., Dekel A., 1989, *ApJ*, 336, L5
- Carlesi E. et al., 2016, *MNRAS*, 458, 900
- Colless M., Saglia R. P., Burstein D., Davies R. L., McMahan R. K., Wegner G., 2001, *MNRAS*, 321, 277
- Dekel A., Bertschinger E., Faber S. M., 1990, *ApJ*, 364, 349
- Dolag K., Komatsu E., Sunyaev R., 2016, *MNRAS*, 463, 1797
- Doumler T., Hoffman Y., Courtois H., Gottlöber S., 2013, *MNRAS*, 430, 888
- Dubois Y., Peirani S., Pichon C., Devriendt J., Gavazzi R., Welker C., Volonteri M., 2016, *MNRAS*, 463, 3948
- Efstathiou G., Eastwood J. W., 1981, *MNRAS*, 194, 503
- Freedman W. L. et al., 2001, *ApJ*, 553, 47
- Ganon G., Hoffman Y., 1993, *ApJ*, 415, L5
- Hahn O., Abel T., 2011, *MNRAS*, 415, 2101
- Hoffman Y., Ribak E., 1991, *ApJ*, 380, L5
- Hoffman Y., Ribak E., 1992, *ApJ*, 384, 448
- Huchra J. P. et al., 2012, *ApJS*, 199, 26
- Jasche J., Wandelt B. D., 2013, *MNRAS*, 432, 894
- Jha S., Riess A. G., Kirshner R. P., 2007, *ApJ*, 659, 122
- Kitaura F. S., Enßlin T. A., 2008, *MNRAS*, 389, 497
- Kitaura F.-S., 2013, *MNRAS*, 429, L84
- Knollmann S. R., Knebe A., 2009, *ApJS*, 182, 608
- Lavaux G., 2010, *MNRAS*, 406, 1007
- Lavaux G., 2016, *MNRAS*, 457, 172
- Lavaux G., Hudson M. J., 2011, *MNRAS*, 416, 2840
- Lavaux G., Mohayaee R., Colombi S., Tully R. B., Bernardeau F., Silk J., 2008, *MNRAS*, 383, 1292
- Lee M. G., Freedman W. L., Madore B. F., 1993, *ApJ*, 417, 553
- Mathewson D. S., Ford V. L., Buchhorn M., 1992, *ApJS*, 81, 413
- Planck Collaboration XVI, 2014, *A&A*, 571, A16
- Sachs R. K., Wolfe A. M., 1967, *ApJ*, 147, 73
- Schaye J. et al., 2015, *MNRAS*, 446, 521
- Shaya E. J., Tully R. B., Hoffman Y., Pomarède D., 2017, *ApJ*, 850, 207
- Sheth R. K., Diaferio A., 2001, *MNRAS*, 322, 901
- Skrutskie M. F. et al., 2006, *AJ*, 131, 1163
- Sorce J. G., 2015, *MNRAS*, 450, 2644
- Sorce J. G., Tempel E., 2017, *MNRAS*, 469, 2859
- Sorce J. G., Tempel E., 2018, *MNRAS*, 476, 4362
- Sorce J. G., Courtois H. M., Gottlöber S., Hoffman Y., Tully R. B., 2014, *MNRAS*, 437, 3586
- Sorce J. G. et al., 2016a, *MNRAS*, 455, 2078
- Sorce J. G., Gottlöber S., Hoffman Y., Yepes G., 2016b, *MNRAS*, 460, 2015
- Sorce J. G., Colless M., Kraan-Korteweg R. C., Gottloeber S., 2017a, *MNRAS*, 471, 3087
- Sorce J. G., Hoffman Y., Gottlöber S., 2017b, *MNRAS*, 468, 1812
- Springob C. M., Masters K. L., Haynes M. P., Giovanelli R., Marinoni C., 2007, *ApJS*, 172, 599
- Struble M. F., Rood H. J., 1988, *Sky Telesc.*, 75, 16
- Tempel E., Kipper R., Tamm A., Gramann M., Einasto M., Sepp T., Tuvikene T., 2016, *A&A*, 588, A14
- Tonry J. L., Dressler A., Blakeslee J. P., Ajhar E. A., Fletcher A. B., Luppino G. A., Metzger M. R., Moore C. B., 2001, *ApJ*, 546, 681
- Tully R. B., 2015, *AJ*, 149, 171
- Tully R. B., Fisher J. R., 1977, *A&A*, 54, 661
- Tully R. B., Shaya E. J., Karachentsev I. D., Courtois H. M., Kocevski D., Rizzi L., Peel A., 2008, *ApJ*, 676, 184
- Tully R. B. et al., 2013, *AJ*, 146, 86
- Tully R. B., Courtois H. M., Sorce J. G., 2016, *AJ*, 152, 50
- van de Weygaert R., Bertschinger E., 1996, *MNRAS*, 281, 84
- Vogelsberger M. et al., 2014, *MNRAS*, 444, 1518
- Wang H., Mo H. J., Yang X., van den Bosch F. C., 2013, *ApJ*, 772, 63
- Wang H. et al., 2016, *ApJ*, 831, 164
- Willick J. A., Courteau S., Faber S. M., Burstein D., Dekel A., Strauss M. A., 1997, *ApJS*, 109, 333
- Zaroubi S., Hoffman Y., Fisher K. B., Lahav O., 1995, *ApJ*, 449, 446
- Zaroubi S., Hoffman Y., Dekel A., 1999, *ApJ*, 520, 413
- Zaroubi S., Bernardi M., da Costa L. N., Hoffman Y., Alonso M. V., Wegner G., Willmer C. N. A., Pellegrini P. S., 2001, *MNRAS*, 326, 375
- Zeldovich Y. B., Sunyaev R. A., 1969, *Ap&SS*, 4, 301

This paper has been typeset from a \LaTeX file prepared by the author.